

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-32422
(P2000-32422A)

(43) 公開日 平成12年1月28日 (2000.1.28)

(51) Int.Cl. ⁷	識別記号	F I	テームト* (参考)
H 0 4 N 7/15		H 0 4 N 7/15	5 C 0 6 4
G 0 6 F 3/16	3 2 0	G 0 6 F 3/16	3 2 0 A
			3 2 0 H

審査請求 有 請求項の数 2 O L (全 6 頁)

(21) 出願番号 特願平10-193145

(22) 出願日 平成10年7月8日 (1998.7.8)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 加村 ゆかり

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100108578

弁理士 高橋 昭男 (外3名)

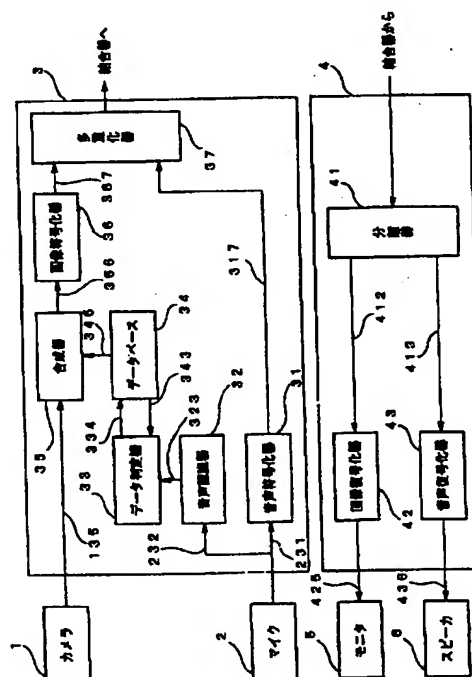
Fターム(参考) 5C064 AA02 AC02 AC06 AC08 AC17
AD02 AD09

(54) 【発明の名称】 画像伝送装置

(57) 【要約】

【課題】 送信する音声を音声認識し、その音声に対応したデータをデータベースより選択し、送信映像と合成して送信する画像伝送装置を提供する。

【解決手段】 マイク2からの音声は音声符号化器31で符号化されて多重化器37へ供給されると同時に、音声認識器32で認識されてデータ判定機33へ供給される。データ判定機33によりこの音声に対応するデータがデータベース34より選択され、合成器35へ供給される。合成器35で前記データがカメラ1からの映像と合成され、画像符号化器36で符号化されて多重化器37へ供給される。多重化器37では画像符号化器36からの映像と音声符号化器31からの音声を多重化する。分離器41では、受信された多重化信号が符号化映像と符号化音声に分離され、符号化映像は画像復号化器42で復号化されてモニタ5へ、符号化音声は音声復号化器43で復号化されてスピーカ6へ出力される。



1

【特許請求の範囲】

【請求項 1】 送信された音声を音声認識する音声認識手段と、
前記音声認識手段の認識結果に対応するデータを蓄積している蓄積手段と、
前記データを前記蓄積手段より選択する選択手段と前記選択手段より選択された前記データと前記送信された画像とを合成する合成手段とを具備し、音声に対応したデータを映像と共に送信することを特徴とする画像伝送装置。

【請求項 2】 前記蓄積手段には、予め手話映像が蓄積されていることを特徴とする請求項 1 記載の画像伝送装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、テレビ会議システムにおいて、音声認識を用いてデータの選択を行い、映像と合成して送信する画像伝送装置に関する。

【0002】

【従来の技術】 一般に、テレビ会議システムを使用する場合、会議中相手の映像を見ていることが多いが、音声だけでは分かりにくいような内容の場合、補足的に映像情報を用いることが会議の運用に有効である。従来のテレビ会議システムにおいては、補足情報として映像情報、例えば図面などをデータとして送って表示させる場合、予め使用するデータを準備しておかなくてはならなかった。

【0003】

【発明が解決しようとする課題】 しかしながら、従来のテレビ会議システムでは、会議の展開に応じた適切な映像情報をタイムリーに使用するのには難しいという問題があった。その理由は、送信するデータは予め用意されたものか、送信されている映像や音声の情報に左右されない内容のものを映像や音声と多重化して送信することしできないからである。

【0004】 本発明はこのような点を考慮してなされたもので、テレビ会議において、送信する音声を生認識により認識し、その音声に対応したデータをデータベースより選択し、送信映像と合成して送信する画像伝送装置を提供することを目的とする。

【0005】

【課題を解決するための手段】 請求項 1 記載の発明は、テレビ会議システムにおいて、送信された音声を生認識する音声認識手段と、前記音声認識手段の認識結果に対応するデータを蓄積している蓄積手段と、前記データを前記蓄積手段より選択する選択手段と前記選択手段より選択されたデータと送信された画像とを合成する合成手段とを具備し、音声に対応したデータを映像と共に送信することを特徴とする画像伝送装置である。

【0006】

2

【発明の実施の形態】 以下、図面を参照して本発明の一実施形態について説明する。図 1 は本発明の一実施形態による画像伝送装置の構成を示すブロック図である。この図において、3 は CODEC 送信部、4 は CODEC 受信部である。CODEC 送信部 3 において、符号 31 はマイク 2 から供給された音声を生符号化する音声符号化器、32 はマイク 2 から供給された音声を生認識する音声認識器、33 は音声認識器 32 で認識された音声に対応するデータをデータベース 34 より選択するデータ判定機、35 はカメラ 1 より供給された映像とデータベース 34 より供給されたデータとを合成する合成器、36 は合成器 35 より供給された映像を生符号化する画像符号化器、37 は音声符号化器 31 から供給された符号化音声を生画像符号化器 36 から供給された符号化映像と多重化する多重化器である。また、CODEC 受信部 4 において、41 は符号化音声と符号化映像の多重化信号を生符号化映像と符号化音声とに分離する分離器、42 は分離器 41 からの符号化映像を生復号化して線 425 を介してモニター 5 へと出力する画像復号化器、そして 43 は分離器 41 からの符号化音声を生復号化して線 436 を介してスピーカ 6 へと出力する音声復号化器である。

【0007】 次に、上記構成による画像伝送装置の動作について説明する。カメラ 1 からの映像は、線 135 を介して合成器 35 に供給される。マイク 2 からの音声は、線 231 を介して音声符号化器 31 へ、また線 232 を介して音声認識器 32 へと供給される。音声符号化器 31 では供給された音声を生符号化し、線 317 を介して多重化器 37 へと供給する。音声認識器 32 では供給された音声を生認識し、そのデータを線 323 を介してデータ判定機 33 へと供給する。データ判定機 33 では認識された音声に対応するデータを線 343 を介してデータベース 34 より選択し、その判定結果を生線 334 を介して再びデータベース 34 へ供給する。データベース 34 は、データ判定器 33 から供給された情報に対応するデータを線 345 を介して合成器 35 へ供給する。合成器 35 ではカメラ 1 より供給された映像とデータベース 34 より供給されたデータとを合成し、線 356 を介して画像符号化器 36 へと供給する。画像符号化器 36 では供給された映像を生符号化し、線 367 を介して多重化器 37 へと供給する。多重化器 37 では画像符号化器 36 から供給された符号化映像と音声符号化器 31 から供給された符号化音声を生多重化し、結合器へ出力する。

【0008】 分離器 41 では結合器を介して供給された多重化信号を生符号化映像と符号化音声とに分離し、符号化映像は線 412 を介して画像復号化器 42 へ供給され、符号化音声は線 413 を介して音声復号化器 43 へと供給される。画像復号化器 42 では符号化映像を生復号化して線 425 を介してモニター 5 へと供給する。モニター 5 では送信された映像を表示する。音声復号化器 43 では符号化音声を生復号化して線 436 を介してスピーカ 6 へと供給す

3

る。

【0009】図2は、複数の画像伝送装置によって構成されるテレビ会議システムの構成を示す図である。複数の装置間におけるデータの送受信は、ISDN (Integrated Service Digital Network) などのネットワークを介して行われる。装置A側のカメラ1からの映像およびマイク2からの音声は、CODEC送信部3に供給される。これらの映像および音声は、CODEC送信部3において前述のような処理を施されて多重化されたデータとして結合器7へ供給される。結合器7へ供給されたデータは、ISDN 8を介して装置B側の結合器7' へ供給される。

【0010】結合器7' に供給されたデータは、装置B側のCODEC受信部4' へ供給され、そこで前述のように、再び映像データと音声データとに分離され、さらに復号化されて、モニタ5' およびスピーカ6' へそれぞれ供給される。以上は装置A側から装置B側へのデータの流れについて述べたが、装置B側から装置A側へのデータの流れについても全く同様であるので、これについての説明は省略する。

【0011】本発明により、データベースに手話の映像データを蓄積しておき、音声認識器から送信された音声データに対応した手話の映像データをデータ判定器がデータベースより選択することによって、耳の不自由な人に手話を知らない人の会話を聞かせることが可能になる。また、データベースに他国語のデータを蓄積しておき、送信する言葉に対応する翻訳語をデータ判定器が選択することによって、異なる言語を話す人同士の会話が可能となる。

【0012】

4

*【発明の効果】以上説明したように、本発明によれば、送信される音声を音声認識手段により認識し、その音声に対応したデータを蓄積手段より選択し、送信された映像と合成して送信するので、送信側の音声に対応したデータを、音声と同時に受信側に表示できるという効果が得られる。

【図面の簡単な説明】

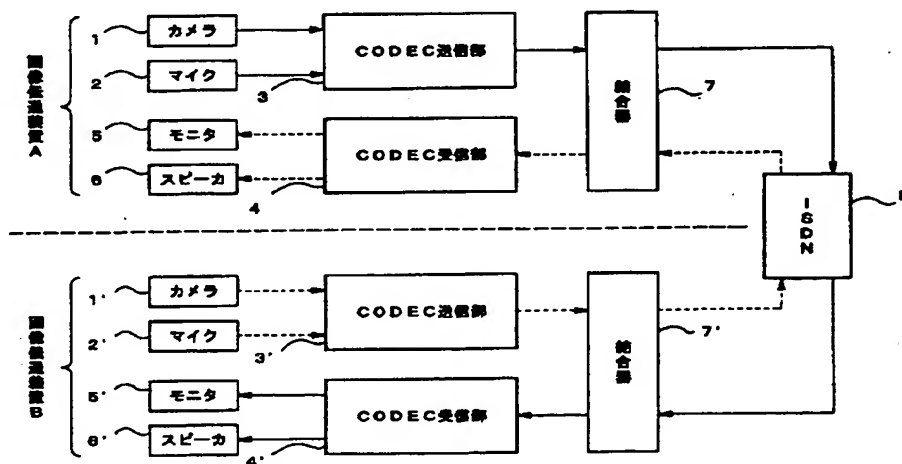
【図1】 本発明の一実施形態による画像伝送装置の構成を示すブロック図である。

【図2】 本発明の一実施形態による、複数の画像伝送装置間でのデータのやりとりを示すブロック図である。

【符号の説明】

- 1, 1' . カメラ
- 2, 2' . マイク
- 3, 3' . CODEC送信部
- 31. 音声符号化器
- 32. 音声認識器
- 33. データ判定機
- 34. データベース
- 35. 合成器
- 36. 画像符号化器
- 37. 多重化器
- 4, 4' . CODEC受信部
- 41. 分離器
- 42. 画像復号化器
- 43. 音声復号化器
- 5, 5' . モニタ
- 6, 6' . スピーカ
- 7, 7' . 結合器
- 8. ISDN

【図2】



【手続補正書】

【提出日】平成10年8月19日(1998. 8. 1

9)

【手続補正1】

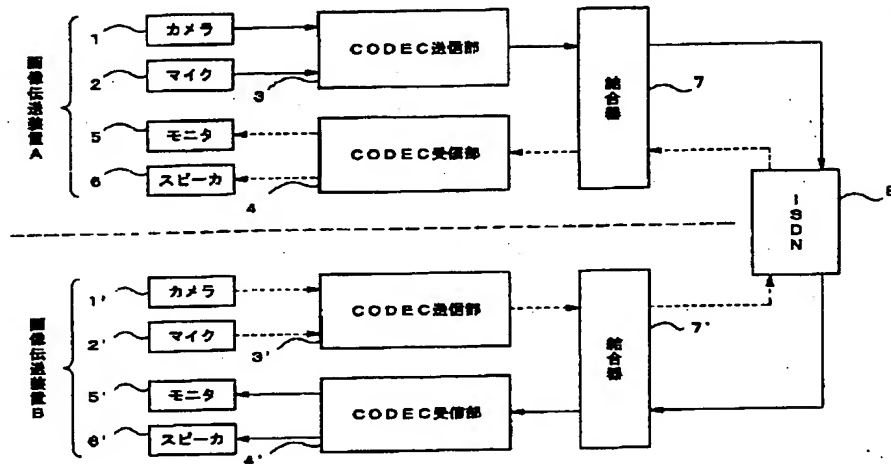
【補正対象書類名】図面

【補正対象項目名】全図

【補正方法】変更

【補正内容】

【図2】



【図1】

